

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-202390

(43)Date of publication of application : 27.07.2001

(51)Int.Cl.

G06F 17/30

(21)Application number : 2000-378217

(71)Applicant : XEROX CORP

(22)Date of filing : 13.12.2000

(72)Inventor : GLANCE NATHALIE S

(30)Priority

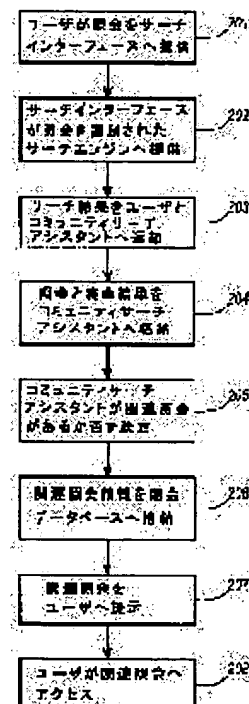
Priority number : 1999 464454 Priority date : 14.12.1999 Priority country : US

(54) NETWORK BASE INFORMATION RETRIEVAL SYSTEM AND DOCUMENTARY SEARCH PROMOTING METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method and a system for promoting (facilitating) the search of information from a network connection information source.

SOLUTION: In a step 204, a community search assistant stores an inquiry and the result and in a step 205, it is determined whether a preceding related inquiry is present or not. The reference of determination on whether the inquiry is related or not depends on the 'overlap' of inquiries. In a step 206, related inquiry information is maintained in an inquiry data base together with the inquiry. Next, in a step 207, the inquiry found to have a relation to a presented inquiry is presented to a user by a search interface. At such a time, in a step 208, the user can access one or more identified related inquiries.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2001-202390
(P2001-202390A)

(43) 公開日 平成13年7月27日 (2001.7.27)

(51) Int.Cl. ⁷	識別記号	F I	テマコード* (参考)
G 0 6 F 17/30	3 4 0 1 1 0	G 0 6 F 17/30	3 4 0 Z 1 1 0 F

審査請求 未請求 請求項の数 2 O L (全 11 頁)

(21) 出願番号 特願2000-378217(P2000-378217)
(22) 出願日 平成12年12月13日 (2000.12.13)
(31) 優先権主張番号 4 6 4 4 5 4
(32) 優先日 平成11年12月14日 (1999.12.14)
(33) 優先権主張国 米国 (US)

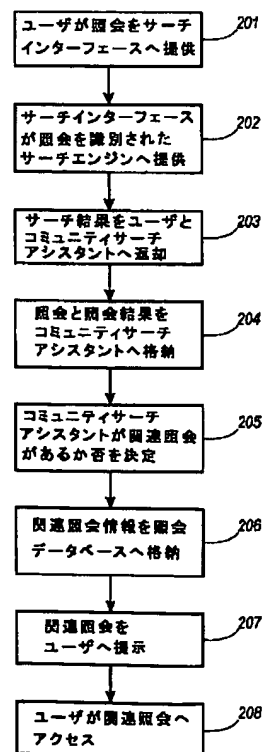
(71) 出願人 590000798
ゼロックス コーポレーション
XEROX CORPORATION
アメリカ合衆国 コネティカット州・スタ
ンフォード・ロング リッチ ロード・
800
(72) 発明者 ナタリー エス. グランズ
フランス国 メイラン 38240 アレド
ウ ラ ブラリ 14
(74) 代理人 100079049
弁理士 中島 淳 (外1名)

(54) 【発明の名称】 ネットワークベース情報検索システム及びドキュメントサーチ促進方法

(57) 【要約】

【課題】 ネットワーク接続情報源から情報をサーチすることを促進 (容易に) するための方法及びシステムを提供すること。

【解決手段】 ステップ204において、コミュニティサーチアシスタントは、照会とその結果を記憶し、ステップ205において、前の関連する照会があるか否かを決定する。照会が関連するか否かの決定の基準は、照会同士間の結果の"オーバーラップ"に依存する。ステップ206において、関連する照会情報は、照会データベース中に照会と共に維持される。次に、ステップ207において、提出された照会に関連することが発見された照会がサーチインターフェースによってユーザへ提示される。この時、ステップ208において、ユーザは、識別された関連照会の一つ又はそれより多くにアクセス出来る。



【特許請求の範囲】

【請求項1】 ネットワークベース情報検索システムであって、

a) ユーザが一つ又はそれより多くのサーチエンジンに照会することを可能とするサーチインターフェースと、
b) 前記ユーザによって提出された照会に関連する照会リストを提供するコミュニティサーチアシスタントと、
を備え、前記コミュニティサーチアシスタントは、更に、

b1) 照会結果リストを含む照会情報を得るために、関連照会を識別するために使用される照会グラフ生成サーチエンジンとのインターフェースするためのサーチエンジンインターフェースと、

b2) 前記対応する照会結果リストに基づいて識別される関連照会が識別出来るように前記照会情報を処理するための照会関係処理手段と、

を備え、

前記システムは、更に、

c) 前記照会情報を格納するための照会データベースを格納するために、前記コミュニティサーチアシスタントに結合された記憶デバイスを備える、ネットワークベース情報検索システム。

【請求項2】 情報検索システムのためにドキュメントサーチを促進する方法であって、

a) 一セットの照会を得るステップと、

b) 前記セットの照会に基づいて関連照会のグラフを生成するステップと、

c) ユーザ照会に応答して、前記関連照会のグラフに基づいて関連照会を識別するステップと、

d) 前記関連照会を前記ユーザに選択のために提示するステップと、を備えるドキュメントサーチ促進方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、情報検索の分野に関し、より詳細には、ワールドワイドウェブのようなネットワークベースの情報システムにおいて情報をサーチする分野に関する。

【0002】

【従来の技術】例えば、ワールドワイドウェブ（以降ウェブと呼ばれる）で発見されるような、電子情報の常に増加する世界において情報のサーチは、満足出来るもの出あると共にフラストレーションを起し得るものである。このような環境においてリンクされたドキュメントの非常に大きなコレクションの意味を取り且つ情報を捜すことは、特殊な助け無しでは困難である。2種類の助けが情報の探索者を助けるために発展されている。その内の第1のものは、分類システムのある形態において情報を要約し且つクラスタリングする構造やツールである。このようなものの例は、ライブラリカードカタログ（library card catalogs）及び

ヤフーウェブサイト（URL: <http://www.yahoo.com>）である。第2のものは、ユーザのニーズに関連する情報を予想し、従って情報の提示を命令しようとするシステムである。これらの例は、リコス（Lycos）（URL: <http://www.lycos.com>）のようなサーチエンジンを含み、これは、情報ニーズのユーザの仕様をワード及びフレーズの形態で取り、ユーザのニーズに関連すると予想されるドキュメントのランク付けリストを戻す。

【0003】ツールの他のクラスは、推奨（recommender）システムである。推奨システムは、幾つかの予め決定されたフィルタ処理基準に基づいて見る価値のある推奨された引き続くウェブページのリストを提供する。一つのそのような推奨ツールは、アレックスインターネットウェブ（Alex Internet Web）サイト（URL: <http://www.alex.com>）で提供される”推奨（Recommend）”特徴である。この”推奨”特徴は、ユーザが現在見ているウェブページに基づいて検索し見ることを望む関連ウェブページのリストを提供する。

【0004】推奨システムの特性は、推奨システムが推奨する場合において先の照会から得られた知識を利用する意味において合作的（共同的）であることである。一般に、共同的サーチ方法は、前の関連照会の結果を利用して探索する。共同的サーチ方法はまだほんの初期である。しかしながら、ウェブにおいて、サーチエンジンは、共同的サーチのために複数の簡単な技術を組み込む事を始めている。ダイレクトヒット（Direct Hit）（www.directhit.com）は、大衆性（popularity）エンジンを確立し、それは、非常に単純な投票（voting）メカニズムを使用した動作する。この大衆性エンジンを使用するサーチエンジンは、ユーザによって入力された照会とそれらのユーザが従うリンクを単純に追跡する。ユーザは、結果を引き続いて見ることに基づいて投票し、将来において、同じ照会の場合、前のユーザの動作を考慮する順序付けを有する結果を生じる。従って、ダイレクトヒットの大衆性エンジンを使用するサーチエンジンに照会を入力すると、その照会に対して最もポピュラーな結果を戻す。ダイレクトヒットは、（ユーザの照会ターム（用語）のサブセット（部分セット）或いはスーパーセット（拡大セット）を使用して）ユーザの照会範囲を拡大したり縮小することによって働く関連サーチ技術を有する。

【0005】

【発明が解決しようとする課題】本発明は、ワールドワイドウェブ（WWW）で見付けられるようなネットワーク接続情報源から情報をサーチすることを促進（容易に）するための方法及びシステムを提供することを課題とする。

【0006】

【課題を解決するための手段】本発明は、照会を生成するウェブユーザの全体の能力を利用する。第1に、所与の時間内にサーチエンジン（単数又は複数）に提出された全ての照会のグラフが構成される。各ノードが照会である。二つの照会が関連していると決定された場合それらの二つのノードの間にリンクが確立される。関連性（関連度）の決定は、照会自体中の実際のターム（用語）ではなくて、照会によって戻されるドキュメントに依存する。例えば、関連性の基準は、各照会に対して戻されるトップ10件のドキュメントに対して、二つのリストが共通に少なくとも一つのドキュメントを有するという事であってもよい。新たな照会が受信されると、関連する照会が識別される。更に、第1の祖先から第2の祖先へ、第3の祖先へ、等一セットの結果までのパスに従う順序付けられた方法でユーザが関連照会のネットワークを追跡する方法を開示している。

【0007】本発明の第1の態様は、ネットワークベース情報検索システムであって、

a) ユーザが一つ又はそれより多くのサーチエンジンに照会することを可能とするサーチインターフェースと、
b) 前記ユーザによって提出された照会に関連する照会リストを提供するコミュニティサーチアシスタントと、
を備え、前記コミュニティサーチアシスタントは、更に、

b1) 照会結果リストを含む照会情報を得るために、関連照会を識別するために使用される照会グラフ生成サーチエンジンとのインターフェースするためのサーチエンジンインターフェースと、

b2) 前記対応する照会結果リストに基づいて識別される関連照会が識別出来るように前記照会情報を処理するための照会関係処理手段と、を備え、前記システムは、更に、

c) 前記照会情報を格納するための照会データベースを格納するために、前記コミュニティサーチアシスタントに結合された記憶デバイスを備える、ネットワークベース情報検索システムである。

【0008】本発明の第2の態様は、情報検索システムのためにドキュメントサーチを促進する方法であって、

a) 一セットの照会を得るステップと、

b) 前記セットの照会に基づいて関連照会のグラフを生成するステップと、

c) ユーザ照会に応答して、前記関連照会のグラフに基づいて関連照会を識別するステップと、

d) 前記関連照会を前記ユーザに選択のために提示するステップと、を備えるドキュメントサーチ促進方法である。

【0009】

【発明の実施の形態】ワールドワイドウェブで発見されるようなネットワークに接続された情報源から情報をサ

ーチすることを促進するための方法及びシステムが開示される。このシステムは、ユーザの照会及び照会の結果に基づく知識をユーザが暗黙のうちに共有する合同メカニズムを提供する。本発明の好適な実施の形態は、ワールドワイドウェブ（以降ウェブと呼ぶ）として知られるインターネットの部分に常駐するリンクドキュメント（例えば、ウェブページ）のコレクションをサーチするための実施される。ウェブは、リンクドキュメントを格納しアクセスするための使用されるインターネットの部分である。ウェブページは、そのユニフォームリソースロケータ（URL）によってアクセスされる。ウェブサイトは、種々のタイプのウェブロケーションに対する総括的用語である。一つの例では、ウェブサイトは、会社、教育機関等のワールドワイドウェブ上にプレゼンスを有するエンティティと関連する関連ウェブページのコレクションである。或いは、ウェブサイトは、ワールドワイドウェブへのエントリポイントをウェブのユーザに提供するための”入口”であってもよい。また、ウェブサイトは、ウェブを介してアクセスされ得るドキュメントのコレクションを含むリポジトリや情報源であってもよい。一般に、ウェブサイトは、高レベルURL（即ち、www.Abccompany.com）を有し、そのサイトのページは、低レベル修飾子（即ち、www.Abccompany.com/page1.html）を有する。

【0010】ウェブブラウザは、ユーザがウェブ上の常駐するドキュメントを走査し見ることを可能とするツールである。ブラウザの関連する他のレンダリング手段は、ドキュメントの音声部分を聞いたり或いはドキュメントのビデオや画像部分を見ることを可能とする。商用的に利用可能なウェブブラウザの例は、ネットスケープコミュニケーションコーポレーション（Netscape Communication Corporation）から利用可能なネットスケープナビゲータ（Netscape Navigator（登録商標））、及びマイクロソフト（登録商標）コーポレーションから利用可能なインターネットエクスプローラ（Internet Explorer）を含む。

【0011】サーチエンジンは、ワールドワイドウェブ上で利用可能なドキュメント（例えば、ウェブページ）をサーチするウェブサイトである。サーチエンジンは、一般的には、サーチを促進するためにワールドワイドウェブ上の他のサイトウェブページを識別する情報をインデックス付けする。利用可能なサーチエンジンの例は、アルタヴィスタ（Alta Vista（登録商標））、ヤフー！（Yahoo!）（登録商標）（www.yahoo.com）、及びエキサイト（エスエム）（Excite(sm)）（www.excite.com）である。

【0012】ウェブ及びサーチエンジンを介してそれに

タッピングする人々の数が指数的に増加し続けている。従って、任意の所与の情報のニーズに対して、ある人が最近においてそれに対して優れた照会の表示を公式化した高い可能性があるという仮説を立てることに合理性がある。本発明では、ユーザの検索能力が類似の情報を検索した他のユーザの知識を捕獲することによって向上される。概念上は、行う事は、所与の期間内で検索エンジン（単数又は複数）に提出された全ての照会のグラフを構築することである。キーとなるアイデアは、照会グラフの構築は、ウェブの単一のユーザ使用（例えば、検索）を共同使用、即ち、全てのユーザが他のユーザによって提出された照会の知識ベースにタップ出来ること、に変換することである。グラフは、照会同士間の関係を指示すると共に種々の形態であり得る。キーとなることは、照会同士間の関係が維持出来且つ容易に決定可能であるある表示がある事である。

【0013】本実施の形態において、関連性の決定は、照会によって戻されるドキュメントに依存するが照会中の実際の用語（ターム）に依存しない。例えば、関連性の基準は、照会毎に戻されるトップ10件のドキュメントに対して、その二つのリストが、共通に少なくとも一つのドキュメントを有する事であってもよい。最後に、本発明の好適な実施の形態は、ユーザが、第1の祖先から、第2の祖先へ、また第3の祖先へのように、望ましいセットの結果までのパスに従う順序付けられた方法で関連する照会のネットワークの提示を追跡可能とする方法を記述する。

【0014】図1は、本発明を実施するための基本的構成要素を示すブロック図である。図1を参照すると、サーチインターフェース101は、ユーザ照会102が一つ若しくはそれより多くの検索エンジン103に送信されるのを可能とする。この検索エンジン103は、ネットワーク106（例えば、インターネット）を介してサーチインターフェース101に連結される。また、ネットワーク106には照会グラフ検索エンジン107が接続される。照会グラフ検索エンジンは、照会グラフを確立し且つその結果を照会データベース105へ提供するために特別に使用されるものである。

【0015】検索エンジンは、ウェブアクセス可能であってもよいし、検索は、プロプライエタリ（所有権を主張できる）データベースに向けられてもよい。サーチインターフェース101が種々の検索エンジンとインターフェースする方法は、当該技術では公知である。例えば、ドグパイル（dogpile）及びアスクワンス（askOnce）のような種々のメタサーチャーが存在する。その場合、単一の照会が複数の検索エンジンに送信される事が可能であり、且つそれらの種々の検索エンジンからの結果が併合される。本実施の形態において、“ラッパーズ（wrappers）”は、種々の検索エンジンとのインターフェースを行うために開発

される。“ラッパーズ”は、一つの照会を取り上げ、サーチエンジンによって許可されるフォーマット及びプロトコルで照会をその検索エンジンへ提示する。

【0016】サーチインターフェース101によって提供される機能性を増強（拡張）するのがコミュニティサーチアシスタント104である。このコミュニティサーチアシスタント104は、特定の検索エンジンから独立しているが、一般的には、その照会が送られたものとは異なる照会グラフ情報の生成において、特定の検索エンジン、即ち、照会グラフ検索エンジン107、を使用する。これは、所定の特性（例えば、戻された結果をランク付けするために使用される方法）を有する検索エンジンがより良好な結果を提供する事を観察する事に基づく事が好ましい。コミュニティサーチアシスタントは、前の照会情報を維持し、且つ照会同士が関連するか否かの決定を行う。関連性を測定するために使用される検索エンジンは、ユーザ照会のセットが得られたものと同じ検索エンジンである必要はない点に留意すべきである。コミュニティサーチアシスタント104に照会データベース105が連結される。照会データベース105は、照会情報を記憶し検索するために使用される。

【0017】また、本発明は、単一の検索エンジンと対話するのみであるように実施され得る事に留意すべきである。この場合、サーチインターフェース101は、複数の検索エンジン103の内の一箇の部分として統合されてもよい。同様に、コミュニティサーチアシスタント104は、検索エンジン103に直接接続されてもよい。また、本発明は、コミュニティサーチアシスタント104が埋め込まれ、且つサーチインターフェース101の機能性の中に含まれるように実施され得る。

【0018】図2は、本発明と対話する点の基本的ステップを記述するフロー図である。第1に、ステップ201において、ユーザが照会をサーチインターフェースに提出する。照会は、照会用語を含み、且つその照会が向けられるべき特定の検索エンジンを指定してもよい。次に、ステップ202において、サーチインターフェースは、照会を識別された検索エンジンへ提出する。ステップ203において、検索結果は、ユーザ及びコミュニティサーチアシスタントへ戻し提示される。検索結果は、例えば、それらの結果が夫々の検索エンジンから提供されるまま正確に、或いはあらゆるオーバーラップしているドキュメントを除去するために併合されて、種々の方法で提示され得る事を理解すべきである。更に、どの結果がコミュニティサーチアシスタントへ提供されるべきかを決定するための“ルール”が決定されなければならない。それは、何がユーザに提示されてもよいし、選択された検索エンジンの一つからの結果であってもよい。また、コミュニティサーチアシスタントのための結果は、ユーザへ提示されるものと異なっても

よい。例えば、ユーザは、全ての結果を要求してもよく、コミュニティサーチアシスタントには、トップ10の(上から10番目までの)結果が提供される。

【0019】次に、ステップ204において、コミュニティサーチアシスタントは、照会とその結果を記憶し、ステップ205において、前の関連する照会があるか否かを決定する。この好適な実施の形態において、照会が関連するか否かの決定の基準は、照会同士間の結果の”オーバーラップ”に依存する。例えば、結果中の共通ドキュメントの所定数はドキュメントが関連すると決定する基準であってもよい。何れにしても、ステップ206において、関連する照会情報は、照会データベース中に照会と共に維持される。次に、ステップ207において、提出された照会に関連することが発見された照会がサーチインターフェースによってユーザへ提示される。この時、ステップ208において、ユーザは、識別された関連照会の一つ又はそれより多くにアクセス出来る。

【0020】本発明の機能の多くは、コミュニティサーチアシスタントに常駐する。コミュニティサーチアシスタントは、そのコミュニティの知識を抽出するために、特定のコミュニティ(例えば、特定のサーチエンジンのユーザ)からの一セットの照会を処理する。図11は、コミュニティサーチアシスタントの基本的な構成要素を示すブロック図である。図11を参照して、コミュニティサーチアシスタント104は、データベースインターフェース151、サーチエンジンインターフェース152及びインターフェースツースearchインターフェース153を備える。更に、照会グラフジェネレータ154は、照会グラフの生成を可能とし且つ関連照会の識別要求を処理する。照会グラフジェネレータの動作は、以下の紹介グラフ発生、関連照会の発見及びユーザ照会に対する応答に関連して記述される。

【0021】本発明の実施では、ある期間にわたる一つサーチエンジン又は一セットのサーチエンジンに対して行われた一セットの照会へのアクセスを必要とする。そのセットの照会は、ウェブと同程度のユーザベースからでも良いし、或いは会社のイントラネット内のユーザベースに制限されてもよい。より広いユーザベースの利点は、所与の情報ニーズに対して、情報ニーズでのオーバーラップの確率が増加することである。より制限されたユーザベースの利点は、焦点がより狭くなり、ノイズ(非関連照会)が少なくなる。

【0022】本発明において、照会グラフは、照会関係が維持されるデータ構造や他のメカニズムと関連する。概念的には、照会グラフは、各ノードが照会であるノードリンク構造を構築するために必要な情報を含む。二つのノードは、照会が関連するものと計算されると、接続される。照会グラフは、ネットワーク上の現在のコンテンツとその現在の使用法(照会の形態における)を最も良く反映するように拡張された照会セットから増分的

或いは周期的に再構築されてもよい。

【0023】照会グラフ中の個々のレコードを記述するデータ構造は、図3に関連して示される。図3を参照して、照会レコード301は、照会を識別するための照会インデックス302、照会で使用される用語をリストアップする照会用語セクション303、その照会と関連する結果を識別する結果セクション304、関連照会及び当該リストが発生されたデート(日付)を識別する関連照会リスト305、及び最後になされた照会のデートを指示する照会デート306を備える。特定の照会が行われた全ての日付や特定の照会が行われた回数のような、他の情報がセーブされてもよい事が理解されるべきである。また、図示されていないが、各関連照会は、”関連照会”基準の基づく共通のドキュメントの数を指示する関連する”カウント”を有する。

【0024】本実施の形態において、所定のサーチエンジンからの結果は、照会グラフを発生するために使用される。それゆえに、その所定のサーチエンジンが検索のために選択されなかった場合でも、その照会はそれに対して発生される。照会グラフは、SQLのような、関連データベースとともに実施される。

【0025】図4は、初期照会セットでデータベースを構築するためのステップを示すフロー図である。図4を参照して、ステップ401において、一セットの照会Qが得られる。このセットの照会は、ある期間にわたって収集されてもよい。これらの照会は、同じサーチエンジンへ提出されたものである事を必要としない点に留意すべきである。次に、ステップ402において、Q中の各照会q毎に、qが照会グラフを開発するためにサーチエンジンに提供される。このサーチエンジンは、オリジナル照会が送信されたのと同じサーチエンジンではない点に留意すべきである。ステップ403において、各照会q毎に、照会qのトップN個の結果は、リストL(q)にセーブされる。ステップ404において、照会qと対応するリストL(q)は、データベースにセーブされる。

【0026】上述のように、本実施の形態において、照会関連性の度合いは、戻されたドキュメントのオーバーラップに依存する。例えば、関連性の一つの尺度は、一つのサーチエンジンによって或いは選択されたサーチエンジン(複数)によって戻されるトップN個中の共通の項目(アイテム)の数である。本実施の形態において、N=10であるが、あらゆる他の数のドキュメントが使用されてもよい。グラフを構築する目的で、関連性の尺度は、2値であると考えられる。即ち、二つの照会が関連するか或いは関連しないかである。

【0027】関連性の他の尺度は、言語解析や引用解析のような、戻された結果の他のより複雑な解析を含んでもよい。

【0028】図5は、データベース中の照会qに対する

関連する照会を発見するためのステップを示すフロー図である。図5を参照すると、ステップ501において、全ての照会 q に対して、 $L'(q')$ と $L(q)$ の交点 $I(q, q')$ が空でないように Q 中の全ての照会 q' を発見する。換言すれば、リスト $L(q)$ 中にある結果が照会 q' に対するリスト $L'(q')$ にもある。次に、ステップ502において、ステップ501で関連すると識別された照会 q' が関連照会のリスト $R(q)$ に追加される。最後に、ステップ503において、照会 q 、その最新の提出日付、生成されたリストと日付、及び各関連する照会に対するマッチ（合致）カウントがデータベースに格納される。同じ照会 q が複数回提出されてもよいし、この照会ヒストリー情報を維持することが有用である。

【0029】関連する照会を決定するための基準の变量は以下の通りである。

(1) 二つの照会は、少なくとも一对のURLがオーバーラップする（正確に一致することの代わりに）時に関連すると定義され、その場合、オーバーラップは、ある閾値を越えることである。

(2) タイトル、サマリー、ドキュメントの検索内容、ドキュメントの内容のサマリー等のような、応答の他のメタ属性によって照会をマッチする（付き合わせる）。

【0030】図6は、関連する照会が識別される、照会を処理するためのステップを示すフロー図である。図6を参照して、ステップ601において、照会 x がコミュニティサーチアシスタントへ送信される。次に、ステップ602において、コミュニティサーチアシスタントは、照会データベースを構築するために使用される照会グラフサーチエンジンへ照会 x を提出する。これは、この照会がそれが元々提出されたものとは異なるサーチエンジンへ送信され得る事を意味することが理解されるべきである。ステップ603において、新たな照会 x に対するトップ N 個の結果が検索され且つ照会データベース中に $L(x)$ として格納される。上述のように、本実施の形態では、 $N=10$ である。次に、ステップ604において、照会 x に対して、 $L'(q')$ と $L(x)$ の交点 $I(x, q')$ が空ではないように Q 中の全ての q' を発見する。換言すれば、リスト $L(x)$ 中にある結果は、照会 q' に対するリスト $L'(q')$ にもある。上述のように、関連性に対する他の基準が使用されてもよい。次に、ステップ605において、ステップ604で関連すると識別された照会 q' が関連照会のリスト $R(x)$ へ追加される。次に、ステップ606において、照会 x とリスト $R(x)$ が照会データベースに格納される。最後に、ステップ607において、関連照会のリスト $R(x)$ が戻される。本実施の形態において、最初に、共通に最も高いカウントによって順序付けられたトップ12個の関連照会が戻される。

【0031】ステップ608において、関連照会のリス

トのユーザへの提示に先立って、オプションのフィルタ処理が実行されてもよい。このようなフィルタ処理は、幾つかのキーワードを含む照会を除去すること、オーバーラップの大きさや他のメトリック（尺度）によって関連サーチをランク付けすることを含み得る。次に、ステップ609において、関連照会の得られたリストがユーザに提示される。このようなリストは、一般的にあるヒューリスティック（例えば、アルファベット順や関連性の強さ）によって順序付けされる。ユーザが関連照会 q'' を選択すると、次に、ステップ610において、その照会 q'' がステップ602から609で記述されたように処理される。最後に、ステップ611において、ステップ602から610は、ユーザが最早関連照会を見ることを望まないようになるまで繰り返される。

【0032】演算努力の大部分は、大きなセットの照会からグラフをブートストラップしており、これは、予め実行される。オンザフライ演算は、図6に示されるように、新たな照会 Q が提出され、新たな照会と全ての既存のセーブされた照会との間の交点の計算が決定される時にのみ必要とされる。この計算は、データベーステーブル（URL、照会）、を生成する事によって最適化され、URL或いは他のドキュメント識別子によってインデックス付けされることが出来る。

【0033】また、照会が前に提出されており、それが照会データベースに含まれておるか否かを決定し、幾つかの処理ステップ（例えば、図6のステップ602乃至606）をバイパスする事が可能である事に留意すべきである。しかしながら、この場合、日付の無い或いは”無効の”照会情報を含む照会データベースとなり得る。

【0034】本実施の形態において実行されないが、”無効の”照会情報を回避するために種々のアプローチが取られ得る。第1のアプローチは、(1) 照会メモリデータベースをコピーし、(2) コピー中の照会、日付、各照会に対するURLを再決定し、(3) そのセットの関連照会を再計算し、(4) 古い照会メモリデータベースを新しいものと置き換える。現状の演算技術に基づき、並列化レベルに依存し、これは、50000個の照会に対して数時間から数日くらいかかる。

【0035】第2のアプローチ：これは、時間的に安価でより増分的である。合同的サーチアシスタントを介して提出された各照会毎に、そのアシスタントによって発生された全ての関連照会をセーブする。このリストをオフラインで定期的に（例えば、1日に一回）処理する。リスト中の全ての照会に対してURLを見つけ、それらから関連照会を再計算する。無効情報の問題がまだあるが、それは、当該ユーザの興味に関連の無い照会に対してのみである。このアプローチは、高価ではないという利点がある。

【0036】第3の提案：ある日付よりも古い全ての日

付を周期的に削除する。例えば、先週からの照会と関連照会の計算のみを保持する。この場合、日付情報をセーブすることが必要である点に留意すべきである。

【0037】”無効の”照会レコードを最小にするために取られるアプローチは、今までの照会レコードを保持することのコスト／利点のトレードオフに依存する。

【0038】図7と8は、本実施の形態におけるユーザインターフェースを示す。このようなユーザインターフェースは、図1のサーチインターフェース101によって”ウィンドウ”中のコンピュータディスプレイ上に提示される。図7を参照して、ユーザは、一つ又はそれより多くの照会用語を照会エリア701に入力する。この照会は、サーチボタン702を”クリックする”ことによって提示される。その照会が何れのサーチエンジン（単数又は複数）に送信されているかを指示するタイトル703が、図7に指示されている。或いは、その照会が何れのサーチエンジン（単数又は複数）に送信されるべきかを指示するチェックボックスが設けられてもよい。図8を参照して、照会結果が戻されると、それらの結果は、結果エリア801に表示される。結果エリア801は、一般に、サーチ用語に対する関連性に基づいて結果の順序付けを行う。しかしながら、質問されるサーチエンジンの数によって、種々の他の技術が使用され得る。更に、関連照会エリア802が表示される。ユーザが照会を入力すると、ユーザには、一セットの関連サーチ、照会が提示される。関連サーチは、関連性の度合によって、最初に最も高い関連度から順序付けされ得る。

【0039】本実施の形態において、共通して最も高いカウントによって順序付けられたトップ12個の関連照会が提示される。関連照会の全リストを表示し見ることが可能とするために、関連照会エリアは一般的にスクロール可能であったり、或いは全体のユーザインターフェースウィンドウがスクロール可能であってもよいことに留意すべきである。関連サーチをクリックした後、ユーザは、関連サーチに対する関連ドキュメントのセット、及び関連サーチの第2のセット（グラフ中のオリジナル照会から2ホップ離れる）が提示される。このように、一連の関連サーチをクリックすることによって、ユーザは、その照会グラフを通るパスに従うことが出来る。ユーザは、ブラウザの”バック（戻り）”ボタンを使用して、或いはそのためにインターフェースによって提供されたリンクをクリックすることによってバックトラック出来る。

【0040】最後に、本実施の形態において、ユーザインターフェースは、ウェブブラウザの制御下で動作するウェブページとして実施される事に留意すべきである。従って、結果リストは、下層URLを有するドキュメントリストとして提示され、その場合、ユーザは、ドキュメントを”クリック”してもよく、次に、そのURLにあるドキュメントを含む新たなウィンドウが提示され

る。

【0041】サーチエンジンに対するユーザインターフェースは、照会同士間の関係を指示するグラフのユーザインターフェースを提供するために拡張され得る。これは、図8の関連照会エリア802で或いは別個のウィンドウとして提供され得る。このようなユーザインターフェースは、図9に示されている。図9を参照して、関連照会は、ツリー表示を使用して提示され、その場合、ルートにオリジナル照会901があり、リーフ902に種々の関連照会が表示される。これらのリーフは、マルチレベル階層形式で表示され、且つ周知のポイント及びクリックインターフェース技術を使用してアクセス可能である。

【0042】本発明の作用が以下の例を介して更に図示される。ここで見られる結果は、インターネット上で発見される情報の絶えざる変化とダイナミックな性質により、必ずしも複製可能である必要は無い事に留意すべきである。名前がリコール（連想）されないモバイルコンピューティング（mobile computing）についての特定のコンファランス（conference）のウェブページをユーザがサーチしていることを想定する。ユーザは、キーワード”モバイルコンピューティングコンファランス”を使用してサーチエンジンへ照会を提出して、数百の応答を得るが、トップ10に或いはトップ50までに成功はいない。更に、ユーザは、求めているものを見つけるためにその照会用語を最もよく偏向する方法を知らないという事を想定する。サーチを”モバイルコンピューティング”に一般化することは役に立たず、サーチエンジンによって提供される提案された関連サーチや”推奨”の何れでもない。

【0043】そこからコンファランスのウェブサイトをサーチする他の誰かがその名前、この場合は、HUC'99（the International Symposium on Handheld and Ubiquitous Computing（ハンドヘルド及びユービキタスコンピューティングに関する国際シンポジウム））を思い出すという事を想定する。オリジナルユーザは、オリジナル照会を使用してそれを発見する見込みがない。理由は、サーチの対象が実際は、シンポジウムであって、コンファランスではなく、且つキーワードは、ハンドヘルドコンピューティングであって、モバイルコンピューティングではないからである。照会中の用語にオーバーラップがない（モバイルコンピューティング対HUC1999）場合、他の誰かがオリジナルユーザ情報ニーズに対応する照会を入力したという事実を利用することは望ましい。

【0044】二つの照会に共通の用語がない場合でも、戻されたドキュメントから導入され得る二つ用語の間に関係があるということがわかる。

【0045】照会”モバイルコンピューティング”の例

に戻って、その照会”ハンドヘルドコンピューティング
コンファレンス”は、多くのものの内の第1の祖先である。第1の祖先は、それらが直接に関連することを意味する。正確な照会は、他のユーザが実際に提出したものに依存する。二つの照会の両方が、それらのトップ10のウェブページ結果中で”UKCでのモバイルとコンテキストアウェアコンピューティング(Mobile and Context-Aware computing at UKC)”を戻し、従ってそれらは、グラフ中で互いに接続される。用語”ハンドヘルド”が成功すると、ユーザは、この方向へうまくこのグラフへ従う。次に、ユーザは、ハンドヘルドコンピィー手リングコンファレンスに対する第1の祖先(従って、オリジナル照会への第2の祖先)としてハンドヘルドコンピューティングが提示される。理由は、両方の照会は、共にウェブページ”プレスリリースビルディングインダストリーに対するビッドコムブリングスハンドヘルドコンピューティング(Press Release-BidCom Brings Handheld Computing to The Building Industry)”を有するからである。ユーザが求めているページは、ある他のユーザが照会”HUC1999”を作るに十分に親切である場合、次のセットの関連照会(ここでは、オリジナル照会に対する第3次祖先)を追跡することによってそれをより迅速に発見出来る。関連照会、照会用語及び関連結果リストを審査するこのプロセスを介して、ユーザは、サーチしている情報を発見出来る。

【0046】本発明の好適な実施の形態が実施されるコンピュータベースのシステムが図10を参照して記述される。このコンピュータベースシステムと関連動作命令(例えば、ソフトウェア)は、本発明を実施するために使用される回路を具体化する。図10を参照して、コンピュータベースシステムは、バス1001を介して連結された複数の構成要素を備える。バス1001は、複数の並列バス(例えば、アドレス、データ及びステータスバス)及び階層構造バス(例えば、プロセッサバス、ローカルバス、及びI/Oバス)を含み得る。いずれにしても、コンピュータシステムは、更に、内部メモリ1003(この内部メモリ1003は、一般的には、ランダムアクセスメモリとリードオンリメモリの組合せである)からバス1001を介して提供された命令を実行するためのプロセッサ1002を備える。プロセッサ1002は、図4乃至6のフロー図のステップによって示される種々の動作を実行するために使用される。このような動作を実行するための命令は、内部メモリ1003から検索される。プロセッサ1002によって実行されるこのような動作は、図3のフロー図及びこれに関連する記述部分で述べられる処理ステップを含む。これらの動作は、一般に、周知のプログラミング技術を使用して、適切なプログラミング言語でコード化された命令の形態

で提供される。プロセッサ1002と内部メモリ1003は、離散的構成要素であってもよいし、或いはアプリケーション専用集積回路(ASIC)チップのような単一の集積デバイスであってもよい。

【0047】また、バス1001へは、英数字入力のためのキーボード1004、データを記憶するための外部記憶装置1005、カーソルを操作するためのカーソル制御デバイス1006、視覚出力を表示するためのディスプレイ1007、及びネットワーク接続部1008が接続される。キーボード1004は、一般的に、標準QWERTYキーボードであるが、キーパッドのようなテレフォンであってもよい。外部記憶装置1005は、固定或いは取り外し可能な磁気或いは光学ディスクドライブであり得る。カーソル制御デバイス1006、例えば、マウスやトラックボール、は、一般的に、幾つかの機能のパフォーマンスがプログラム可能な、関連するボタンやスイッチを有する。ネットワーク接続部1008は、例えば、対応するソフトウェアを有するローカルエリアネットワーク(LAN)カードやモデムカードのような、ネットワークに接続するための手段を提供する。ネットワークは、最後にインターネットへ接続されているけれども、それは、アメリカンオンライン(登録商標)(American On-Line)、プロディジー(登録商標)(Prodigy)やコンピュサーブ(登録商標)(CompuServe)のようなプロキシサービスや中間ネットワークやダイヤルアップサービスを介してもよい。

【図面の簡単な説明】

【図1】本発明の本実施の形態の基本的構成要素を示すブロック図である。

【図2】本発明の本実施の形態とインターフェースするための基本的ステップを示すフロー図である。

【図3】本発明の本実施の形態におけるデータベースで利用され得るように照会データレコードに対するデータ構造のブロック図である。

【図4】本発明の本実施の形態におけるデータベースで利用され得るように設定される初期照会情報を確立するステップを記述するフロー図である。

【図5】本発明の本実施の形態におけるデータベースで利用され得るように設定される初期照会情報に対する関連照会情報を提供するステップを記述するフロー図である。

【図6】本発明の本実施の形態において利用され得るように新たな照会を処理し且つ引き続いて関連する照会を発見するためのステップを記述するフロー図である。

【図7】本発明の本実施の形態を使用するためにユーザインターフェースを示す図である。

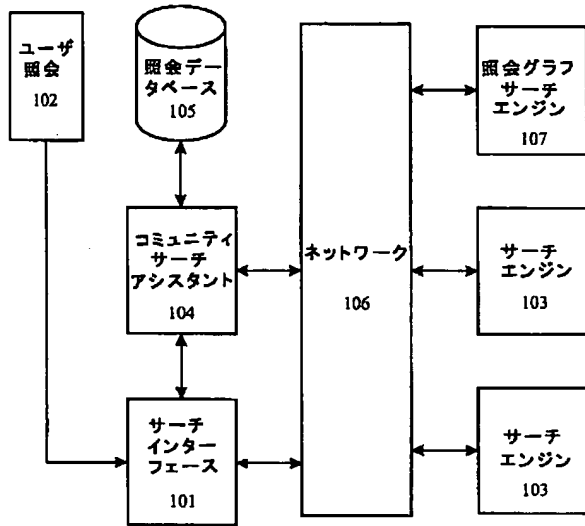
【図8】本発明の本実施の形態を使用するためにユーザインターフェースを示す図である。

【図9】本発明の本実施の形態において利用され得るよ

うにユーザインターフェースに関連照会を表示するためのツリー構造の他の実施の形態の図である。

【図10】本発明の本実施の形態のシステムを実施するのに利用される得るコンピュータシステムのブロック図

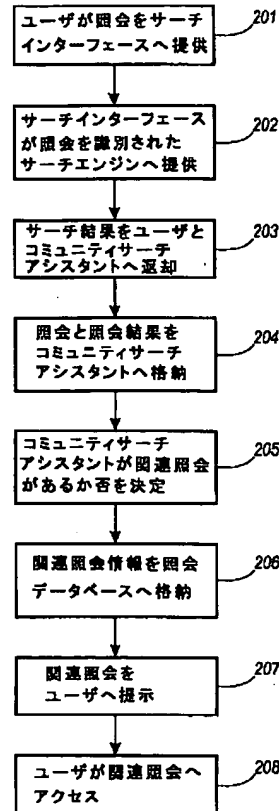
【図1】



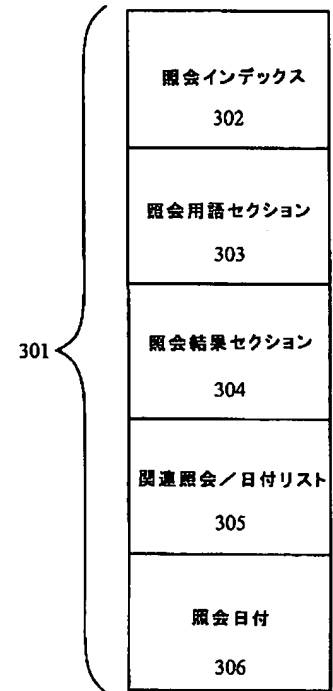
である。

【図11】本発明の本実施の形態において利用され得るようにコミュニティサーチアシスタントの機能構成要素を示すブロック図である。

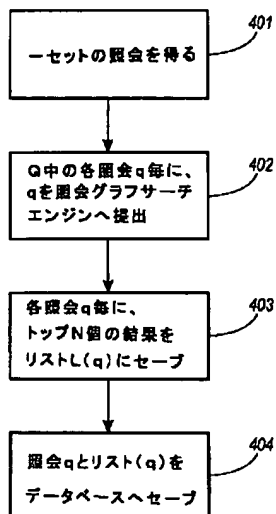
【図2】



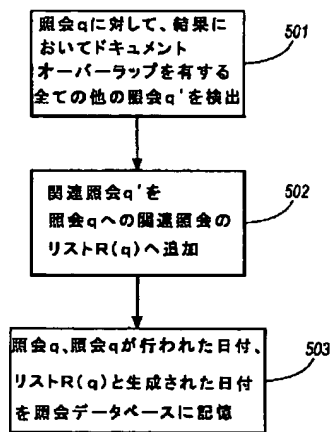
【図3】



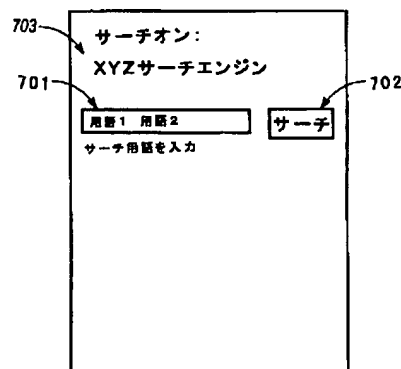
【図4】



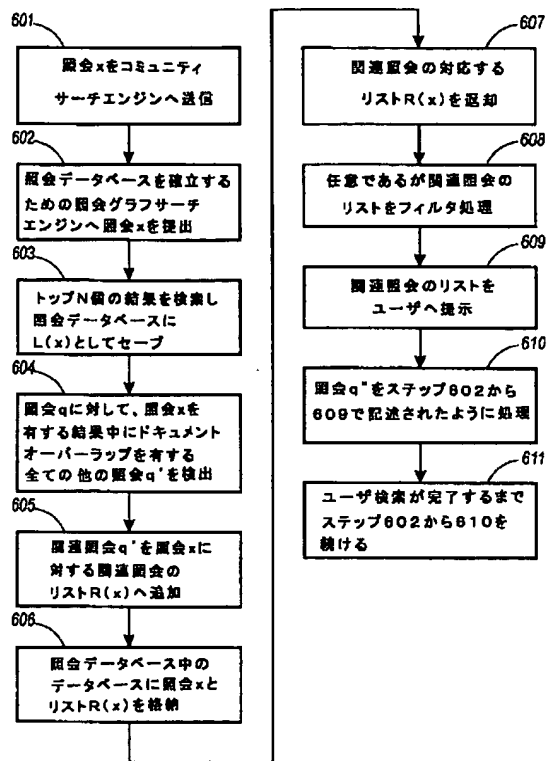
【図5】



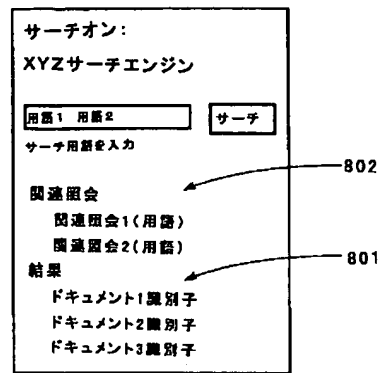
【図7】



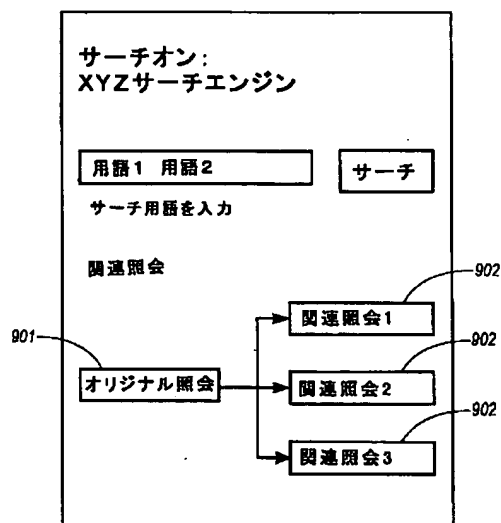
【図6】



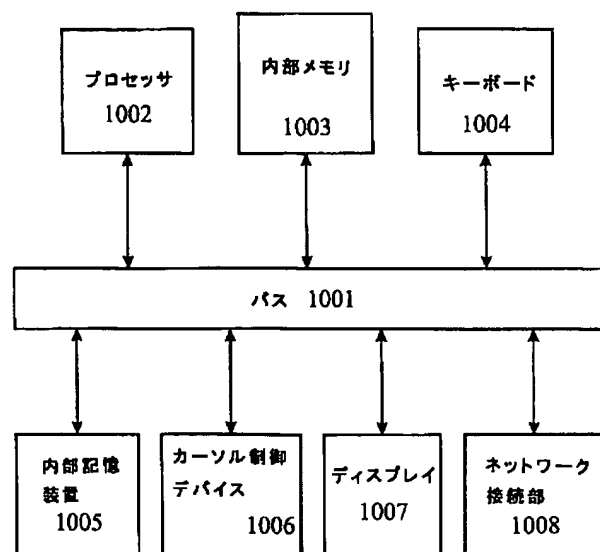
【図8】



【図9】



【図10】



【 11 】

データベースインターフェース 151	104
サーチエンジンインターフェース 152	
インターフェースツ サーチインターフェース 153	
照会グラフジェネレータ 154	